

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

-----X
LOCAL 3621, EMS OFFICERS UNION, DC-37,
AFSCME, AFL-CIO *et al.*,

Plaintiffs,

-against-

THE CITY OF NEW YORK *et al.*,

Defendants.

-----X

JENNIFER E. WILLIS, United States Magistrate Judge:

ORDER

18-cv-4476 (LJL) (JW)

I. BACKGROUND

In December 2023, this Court issued an Order resolving various discovery disputes. Dkt. No. 505. On January 25th, the Plaintiffs filed a letter motion alleging extreme deficiencies in the Defendants' data production and requesting the Court direct the Defendants to permit their data experts to meet and confer with the Plaintiffs' data experts. Dkt. No. 514. Defendants opposed and countered that they had already provided the data they were ordered to produce. Dkt. No. 516. The Parties then submitted additional letters on discovery. Dkt. Nos. 518, 525.

In February, the Court reiterated that the City has been repeatedly directed to provide this data: "after years of acrimony, dozens of conferences, and hundreds of thousands of dollars in sanctions, enough is enough. The demographic data must be turned over for this case to move forward." Dkt. No. 528 (citing Dkt. No. 505). In that Order, this Court warned that "if the Court is dissatisfied with the instructions Defendants provide for determining the demographic and rank information for each individual listed, the Court will consider directing the Defendants to extract and compile the data into a readily useable form proposed by the Plaintiffs, or

alternatively, sanctions.” Id. The Court further warned that if by the time of the April discovery conference, “the demographic data and rank data issues remain unresolved, the Defendants must 1) bring their data expert to the April 3rd conference to answer questions from the Court and 2) be prepared to adequately explain why they have not violated the Court’s previous Order, or the Court will consider imposing sanctions.” Dkt. No. 528.

Pursuant to the February Order, the Parties submitted a joint letter detailing the status of discovery on March 15th. Dkt. No. 529. Defendants argued that they produced the “rank, race, and gender data, the primary focus of the Court’s December 2023 Order...in their entirety.” Dkt. No. 529 at 1. Attached to that submission were various letters the City provided to the Plaintiffs detailing how the data could be analyzed and organized by Plaintiffs’ experts. Dkt. Nos. 529-1, 529-2, 529-3. Defendants’ letters also sought to explain various data discrepancies. Id.

Plaintiffs countered that “the Court has ordered not less than 10 times that Defendant produce the missing Demographic Data.” Dkt. No. 529 (citing Dkt. Nos. 240, 246, 267, 315, 353, 362, 372, 465, 505, and 528). Further, Plaintiffs argued, “Defendants have failed to make a complete production of any of the 5 topics ordered produced by this Court...” Dkt. No. 529 (citing Dkt. No. 465). Plaintiffs claimed they repeatedly requested that the City use specific “customary channels to produce this data such as making a NYCAPs service request or utilizing the FDNY’s Bureau of Technology Development and Services...and [Plaintiffs undertook] repeated efforts to meet and confer with Kamaldeep Deol, a computer systems manager in BTDS with

the sufficient expertise...to timely produce outstanding discovery...Defendants flatly refused to do so...” Dkt. No. 529.

Plaintiffs also submitted a letter from data expert Dr. Shane Thompson which explained that production of the data has been “fraught with significant limitations and errors...critical issues persist with the data production that will continue to impede use of the data.” Dkt. No. 529-4 at 2. Dr. Thompson said that the FDNY “maintains a data warehouse of all FDNY data in the Bureau of Technology Development and Systems (BTDS).” Dkt. No. 529-4 at 6. Dr. Thompson asserted that FDNY experts “write code and pull data from both the data warehouse as well as other databases available to them.” Dkt. No. 529-4 at 6. Dr. Thompson also contended that it would be “much more effective and time efficient to engage BTDS” rather than the current “piecemeal approach” which results in delay and “inconsistencies in the data.” Dkt. No. 529-4 at 5-6.

Hence, Plaintiffs now assert that “severe sanctions are warranted due to Defendants’ repeated noncompliance...the only appropriate remedy...is to impose sanctions in the form of drawing an adverse inference regarding the data that has not been produced.” Dkt. No. 529.

II. LEGAL STANDARDS

A. City Obligated to Produce Items Within its Control

As was discussed in this Court’s December decision, if the requested data exists and is within the City’s “possession, custody, or control” it must be produced. See Fed. R. Civ. P. 34(a)(1); [In re Liverpool Ltd. Partnership](#), No. 21-MC-392 (AJN), 2021 WL

5605044 (S.D.N.Y. Nov. 24, 2021). “Control has been construed broadly by the courts” and has been found to encompass control “by a parent corporation over documents held by its subsidiary...” See [S.E.C. v. Credit Bancorp, Ltd.](#), 194 F.R.D. 469, 472 (S.D.N.Y. 2000)(Sweet, J.)(citing [Cooper Indus. v. British Aerospace Inc.](#), 102 F.R.D. 918, 919 (S.D.N.Y. 1984))(internal citations and quotation marks omitted). Because the City and its own agencies are similarly affiliated, the City is under an obligation to produce responsive documents and data possessed by its own agencies.

B. City Not Obligated to Create Documents But Reports Must Be Created If Necessary for the Data to be Produced in a Useable Form

On the one hand, “Rule 34 cannot be used to compel a party to create a document solely for its production.” [Deng v. New York State Off. of Mental Health](#), No. 13-CV-6801(ALC)(RLE), 2016 WL 11699675, at *3 (S.D.N.Y. Feb. 23, 2016) (quoting [A & R Body Specialty & Collision Works, Inc. v. Progressive Cas. Ins. Co.](#), No. 3:07-CV-929 (WWE), 2014 WL 4437684, at *3 (D. Conn. Sept. 9, 2014)(citing MOORE'S FEDERAL PRACTICE § 30.12[2] (3d ed. 2014))). On the other hand, “where such information is deemed relevant for the purposes of discovery but does not appear to exist in a format readily providable to the plaintiff, requiring creation of such a document or compilation is proper.” [Ford v. Rensselaer Polytechnic Inst.](#), No. 1:20-CV-470(DNH)(CFH), 2022 WL 715779, at *6 (N.D.N.Y. Mar. 10, 2022); see also [In re Domestic Airline Travel Antitrust Litig.](#), No. MC 15-1404 (CKK), 2018 WL 6619855, at *5 (D.D.C. Nov. 27, 2018).

C. Querying Databases & Providing Results Does Not Amount to the Creation of New Documents

“Case law makes clear that simply searching a database and providing the results therefrom” in a report, “does not require creation of a new document beyond the scope of Rule 34.” [Humphrey v. LeBlanc](#), No. CV 20-233-JWD-SDJ, 2021 WL 3560842, at *3 (M.D. La. Aug. 11, 2021) (citing [Apple Inc. v. Samsung Elecs. Co. Ltd.](#), No. 12-630, 2013 WL 4426512, at *3 (N.D. Cal. Aug. 14, 2013); [Bean v. John Wiley & Sons Inc.](#), No. 11-8028, 2012 WL 129809, at *1 (D. Ariz. Jan. 17, 2012); [In re eBay Seller Antitrust Litigation](#), No. 07-1882, 2009 WL 3613511, at *2 (N.D. Cal. Oct. 28, 2009)).

While “a party should not be required to create completely new documents, **that is not the same as requiring a party to query an existing dynamic database for relevant information.**” [N. Shore-Long Island Jewish Health Sys., Inc. v. MultiPlan, Inc.](#), 325 F.R.D. 36, 51 (E.D.N.Y. 2018)(emphasis added)(quoting [Apple Inc. v. Samsung Elecs. Co. Ltd.](#), 2013 WL 4426512, at *3). Indeed, courts “regularly require parties to produce reports from dynamic databases, holding that the technical burden... of creating a new dataset for litigation does not excuse production.” *Id.* at 51 (cleaned up).

The Sedona Principles¹ on Databases explicitly contemplate that when a large portion of a data set is not responsive, custom queries may need to be used to obtain

¹ The Sedona Conference is a “nonprofit legal policy research and education organization” that has a “working group comprised of judges, attorneys, and electronic discovery experts dedicated to resolving electronic document production issues.” See [Aguilar v. Immigr. & Customs Enf't Div. of U.S. Dep't of Homeland Sec.](#), 255 F.R.D. 350, 355 (S.D.N.Y. 2008). Courts have “found the Sedona Principles instructive with respect to electronic discovery issues.” *Id.* (citing [Autotech Techs. Ltd. P'ship v. Automationdirect.com, Inc.](#), 248 F.R.D. 556 (N.D. Ill. 2008) and [Williams v. Sprint/United Mgmt. Co.](#), 230 F.R.D. 640 (D. Kan. 2005)).

the data. See [The Sedona Conference Database Principles Addressing the Preservation and Production of Databases](#), 15 Sedona Conf. J. 171, 181 (2014). As the Sedona Principles explain: “other than situations where a large portion of a given database is responsive, it may be best practice to collect that responsive data by saving a copy of a subset of the database information to a separate location, such as a specifically-designed table, a separate database, or a text delimited file by means of a query or report. In some cases, a pre-existing (‘canned’) query or report may exist that can be used for this purpose. **In other cases, a custom-created query or report will need to be used.**” *Id.* (emphasis added).

D. Reasonable Measures Must be Taken to Ensure Accuracy of Data

Rule 34 not only requires data to be produced in a reasonably useable form, but the data must also be verifiably accurate. See [Chen-Oster v. Goldman, Sachs & Co.](#), 285 F.R.D. 294, 306 (S.D.N.Y. 2012) (citing [The Sedona Conference Database Principles Addressing the Preservation and Production of Databases](#),² March 2011 Public Comment Version, at 32); [DR Distributors, LLC v. 21 Century Smoking, Inc.](#), 513 F. Supp. 3d 839, 936 (N.D. Ill. 2021); [Marin v. Apple-Metro, Inc.](#), No. 12-CV-5274(ENV)(CLP), 2023 WL 2060133 at *8, n.20. (E.D.N.Y. Feb. 8, 2023); [Waskul v. Washtenaw County Cmty. Mental Health](#), 569 F. Supp. 3d 626, 636 (E.D. Mich. 2021). A responding party “must use *reasonable* measures to ensure completeness and

² The Sedona Conference has issued several publications. Some are specific to large databases and others relate to ESI more generally. When the Sedona Principles are referenced here, the specific version featuring the quote is linked. When a court decision cited here references the Sedona Principles, the version referenced in that decision is linked, even if a newer version involving similar language has since been published.

accuracy of the data acquisition.” Chen-Oster, 285 F.R.D. at 306 (emphasis in original)(internal citation and quotation marks omitted).

The Rules “only require a reasonable inquiry” to ensure accuracy. See Diego Zambrano *et. al.*, [Vulnerabilities in Discovery Tech](#), 35 Harv. J.L. & Tech. 581, 592 (2022)(citing Fed. R. Civ. P. 26(g)). **The “standard for the production of ESI is not perfection.”** Chen-Oster, 285 F.R.D. at 306. Rather, the “inquiry must ensure a reasonable degree of accuracy and completeness without being unduly costly. Moreover, the Rules only sanction parties who intentionally or negligently fail to produce relevant documents.” [Vulnerabilities in Discovery Tech](#), 35 Harv. J.L. & Tech. at 592.

The Sedona Principles on Databases also include a specific discussion on validating data queried from large databases. They explain that a “responding party should use reasonable measures to validate that its collection from the database is both reasonably complete and did not inadvertently modify the ESI.” [The Sedona Principles on Databases](#), 15 Sedona Conf. J. at 210. Those principles explained that to “reduce the risk that information extracted from databases contains transcription errors, a responding party that is extracting data from a database and formatting it into a report or file for the purpose of responding to a discovery request should test the proposed dataset to confirm that it includes all expected content and complies with the target format.” Id.

As the most recent version of the Sedona Principles explains: the “touchstone under Rule 34(b)(2)(E), is that a requesting party is entitled to the production of ESI

as it is ordinarily maintained or in a form that is reasonably usable for purposes of efficiently prosecuting or defending the claims and defenses involved in the matter.” [The Sedona Conference, The Sedona Principles, Third Ed. 19 Sedona Conf. J. 1, 175 \(2018\)](#)(cleaned up). Thus a producing party is required to “provide the requesting party a functionally adequate ability to **access, cull, analyze, search, and display the ESI**, as may be appropriate given its nature and the proportional needs of the case.” *Id.* (emphasis added).

III. DISCUSSION

A. The City’s Extensive Efforts to Provide Data

While it seems it took the threat of additional sanctions to push the City to do so, this Court is convinced that the Defendants have now diligently taken steps to produce the data. Reviewing the exhibits and cover letters submitted by the Defendants, this Court finds that on the whole, the City’s production provides Plaintiffs “a functionally adequate ability to access, cull, analyze, search, and display” the data. *See* [The Sedona Conference, The Sedona Principles, Third Ed. 19 Sedona Conf. J. at 175](#) . For this reason, no sanction or adverse inference is appropriate at this time. *See* Fed. R. Civ. P. 37(a)(5)(A)(iii).

B. Issues with the Data Production

For a large database like the City’s, it was appropriate for the Parties to “test[] the proposed dataset” to ensure the “collection from the database is both reasonably complete and did not inadvertently modify the ESI.” *See* [The Sedona Conference Best Practices Commentary on the Use of Search & Information Retrieval Methods in E-](#)

Discovery, 15 Sedona Conf. J. at 210. Having reviewed the results of these validation measures, despite Defendants’ admirable efforts, the Court agrees with Plaintiffs that there were a few issues to be addressed in Defendants’ production. The Court will discuss each in turn.

1. Rank Information

Plaintiffs noted “inconsistencies” in the Rank Data prepared by the City’s data expert Ms. Mason. Dkt. No. 540 at 12. Defendants informed the Court that “Ms. Mason is actually currently working on remedying that defect...” Dkt. No. 540 at 49. Given that Ms. Mason has indicated the issue can be fixed, the Court will not issue any specific order on the Rank Information. The Parties shall provide a status update and proposed timeline on the production of the Rank Data in a joint letter due **May 17, 2024**.

2. Evaluation Forms

In their March letter, the Defendants indicated they would provide evaluation data for 2020-2023, as the performance evaluations from 2004 to 2019 are stored “only in paper form.” Dkt. No. 529 at 2.

However, at the April conference, the Plaintiffs claimed that “we know, based on our clients, that the evaluations are done locally at stationhouses, but then they are PDF’d and emailed through their server to the operations division of that respective member and then to operations human resources. So they do have in one localized place...electronic copies.” Dkt. No. 540 at 23. Plaintiffs said that the “data can easily be queried... there’s a NYCAPS application where people upload their

evaluations directly to operations...” (Dkt. No. 540 at 30, 33). Defendants disagreed, saying “there’s no repository for these evaluations.” Id. at 26-27.

The City is directed to conduct internal questioning to determine if the data is captured electronically as Plaintiffs attest. The City shall provide a status update on the evaluation data by **May 17, 2024**.

3. Leave Time Start Dates and End Dates

In May 2023, the Court ordered the City to produce “information...requested in Plaintiffs’ Document Requests #2, #3, #9, #11, and #12.” See Dkt. No. 465 (citing Dkt. No. 459). Request 3 (m) seeks “start date(s) and end date(s) of any leave that year (e.g. FMLA, etc.) and reason(s) for such leave.” See Dkt. No. 459-1 at 7.

The Court agrees that the cover letter to the production explaining how to calculate such leave time in Dkt. No. 537-5 does not provide Plaintiffs “a functionally adequate ability to access, cull, analyze, search, and display” the data. See [The Sedona Conference, The Sedona Principles, Third Ed. 19 Sedona Conf. J. at 175](#).

At the conference, the Court explained to the City, “What they want is start date and end date of a particular leave...there should be a way that you all can code to query your own system to get that answer... that’s not creating data; that is querying an existing database to get a production.” Dkt. No. 540 at 57-58. The Defendant responded: “I can certainly inquire about this. I don’t know – I don’t know enough about the FISA system to speak to it, and we do have a data person here, but she is not a FISA data person.” Id. at 58-59.

Therefore, the City is directed to produce an expert from FISA-OPA to have a short conference with the Plaintiff's expert to discuss how best to complete the leave data production and whether a custom query would reduce the burden of providing the data. The Parties shall hold the expert conference and provide a status update along with a proposed timeline for the production by **May 17, 2024**.

4. 2010-2014 Missing Data and Missing EMT Promotions and Paramedic Promotions

Defendants data agent, Mr. Manna, told Plaintiffs' data expert that the reason for the missing entries from 2010-2014 was that "no new union contracts occurred during that time, and individuals who reached maximum salaries under the extant agreement wouldn't show up until a new contract was negotiated." Dkt. No. 529-4 at 3.

Plaintiffs' expert, Dr. Thompson, claimed that "this does not explain the absence of data for which salary is only one component...one should be able to identify every year for which an employee is active in the system, regardless of whether union contracts are renegotiated." Id. at 3.

Plaintiffs also argue that "there's no promotional information for EMT to paramedic or to lieutenant...for the entire 20-year period there's no promotional information from EMT to paramedic." Dkt. No. 540 at 70. At the conference, the City's attorney Mr. Olert said, "we explained to plaintiffs' counsel how to see the promotions from EMS to paramedic and from paramedic to lieutenant. However, my understanding is that Ms. Mason can also add that information to the spreadsheet

that we've already produced so that it's in a more clear form." Id. at 86. Ms. Mason cautioned, however, that "we can do that with the caveat that our system is something we keep inhouse because we would like to have access to this information. However, I have to reiterate that source data for the City in this case would be NYCAPS...the NYCAPS data is the gold standard." Id. at 86-87. At the conference, the City's attorney said that for the "2010 to 2014 information...my contact has been at the fire department...that is a person that I can identify." Dkt. No. 540 at 95.

Therefore, the City is directed to produce an expert from DCAS or the fire department to have a short conference with the Plaintiff's expert to discuss how best to complete the promotional data production, including the missing promotions from 2010-2014. The Parties shall hold the expert conference and provide a status update along with a proposed timeline for the production by **May 17, 2024**.

5. Defendants Lack of Production on Promotions to Firefighter

One aspect of the promotional data that Defendants did not produce was information related to promotions from EMT to firefighter. Dkt. No. 540 at 70. Defendants maintain that "firefighter is not an EMS title." Id. Plaintiffs responded that "the order requires that they produce information on all promotions not to EMS but just all promotions of this class." Id. at 72.

In May 2023, the Court ordered the City to produce "information regarding applicants, postings, demographic data, and other relevant information regarding the SEMSS positions... requested in Plaintiffs' Document Requests #2, #3, #9, #11, and #12." See Dkt. No. 465 (citing Dkt. No. 459).

Request 2 sought:

For each job posting or requisition: (1) **relating to an EMS position at the rank of Lieutenant, Captain, Deputy Chief, or Division Commander** and; (2) where one or more applicants were offered a promotion to any such position in any year from 1996 to present, please **produce a complete list of the name, race and gender** of 1) all employees who submitted an application for such job posting or requisition regardless of whether such applicant ultimately completed any further steps in such Promotional Process along with the dates they submitted those applications for each time they applied, 2) all employees who were disqualified from completing the application process at any point in the process either before or after interviewing, and the reason for such disqualification for each time they were disqualified, 3) all employees who were promoted with the date of promotion and title promoted to for each time they were promoted. (emphasis added).

Request 11 said:

Please produce copies of **all promotional postings** and/or announcements issued to publicize each promotional opportunity from 1996 to present, excluding any year for which this information has already been produced, and for each past exam/promotional process produce copies of all documents that reflect the title of the promotional opportunity, copies of all examination announcements, and when they were issued and how they were disseminated and to whom; copies of all application materials, forms, and associated instructions; copies of all examination questions, including any orally asked; copies of all grading guidelines, rating scales, or grading rubrics; copies of all ratings of each candidate, including any initial or interim ratings as well as final ratings; name, job title, and qualifications of all persons who rated or graded any part of the exam or evaluated any of the application materials or forms; grades for each candidate, identified by name and ID number; **names, job titles, and qualifications of the people whom the FDNY relied on to identify job knowledge and other job requirements for each EMS**

promotional exam; and the names and job titles of all DCAS employees, consultants, or other persons with testing expertise who helped develop, administer, or grade the promotional process. (emphasis added).


Nowhere does this explicitly seek a production for the title of “firefighter”. Without a clear reason to do so, the Court declines to include additional titles in an already challenging data production. This request is DENIED.

7. Remaining Data Issues

Plaintiffs raise various other data issues. In his letter, Dr. Thompson highlighted that the “match rate between Employee IDs in prior productions and Employee IDs in new productions is roughly 85%.” Dkt. No. 529-4 at 3. He asserted that “what makes this low match rate most alarming is that it is based on merging only Employee IDs – the field that would seemingly be the most reliable and unchanging.” Dkt. No. 529-4. Defendants have not adequately explained why subsequent data pulls result in such different results.

Additionally, Plaintiffs submitted a letter on April 3rd detailing that Defendants’ use of “canned parameters” impacts “all” recently produced data sets. Dkt. No. 539. That letter included additional assertions of data errors. Id.

The Court disagrees. On balance, the City’s production provides Plaintiffs “a functionally adequate ability to access, cull, analyze, search, and display” the data.

See [The Sedona Conference, The Sedona Principles, Third Ed. 19 Sedona Conf. J. at 175.](#) 

C. Parameters for the Conferences of Data Experts

This is not the first time that a Court in this District has ordered a “committee” of experts to meet to resolve City data production issues. See Winfield v. City of New York, No. 15-CV-05236(LTS)(KHP), 2018 WL 840085, at *3 (S.D.N.Y. Feb. 12, 2018)(“The parties discussed with the Court and now have consented to a novel procedure—a “committee” deposition involving multiple witnesses—which will be outlined in greater detail below.”). The Court believes this resolution will be similarly appropriate here.

On the one hand, the Court agrees with Plaintiffs that the production of this data has been repeatedly ordered and that securing such basic data on City employees should not be so difficult. Indeed, as the Plaintiffs note, the City routinely prepares similar information in accordance with Local Law 18 of 2019, the annual Mayor’s Management Report, and numerous other reports published on a regular basis by NYC Open Data. Dkt. No. 536 at 3, n.1.

On the other hand, however, this Court has serious concerns about the burden of securing data from across different agencies. Accordingly, the Court is interested in limiting the total number of City employees involved in Plaintiffs’ proposed expert conferences and the time commitment such conferences would impose on the City. Despite Plaintiffs’ (welcome) representations that they are willing to merge different data productions and that if the data ultimately does not exist in a useable form that the Plaintiffs would accept that outcome, the Court worries that perhaps no data production would ever satisfy Plaintiffs.

In sum, the Court fears a never-ending, iterative process where Defendants make piece-meal data productions, and then the Plaintiffs constantly seek to repeat the entire process in an attempt to secure new, more favorable data.

Finally, the Court also notes that Judge Liman held a conference on April 23rd in which he reiterated that “this case has been pending for a long time” and that the Parties should “figure out ways to prioritize the discovery that is truly necessary and bring this case to a head.” See Transcript, (April 23, 2024) at 48.

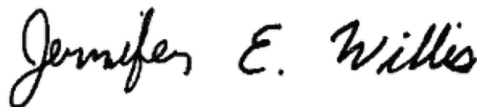
Thus, the Parties are directed to bear these concerns in mind as they move forward with the expert conferences and future data productions.

III. CONCLUSION

In sum, the Court orders the Parties to direct their data experts to meet for short conferences to work cooperatively on the **specific issues identified above** to determine whether custom queries or other methods may lead to a faster, less burdensome production of the data. **By May 17th**, the Parties should complete the expert conferences and provide the Court with a status update and proposed timeline for completing the ordered data productions.

SO ORDERED.

DATED: New York, New York
April 26, 2024



JENNIFER E. WILLIS
United States Magistrate Judge